

# The FHWA Travel Model Improvement Program Workshop over the Web

## The Travel Model Development Series: Part I – Travel Model Estimation

*presented by*  
Thomas Rossi  
Cambridge Systematics, Inc.

October 16, 2008

1

### **Key Message: Purpose of the Webinar Series**

#### **Details:**

Welcome to the FHWA TMIP Workshop over the Web. This workshop is targeted at Transportation modelers who have a low to moderate level of familiarity with the estimation and validation of travel models.

This series of webinars will introduce the development of model estimation data sets, the structures of the various model components, and the procedures for estimating models. The workshop will include lectures, discussion, and “homework,” that participants will be expected to complete between sessions.

This first session will be an introduction to the entire series covering definitions of various model types used in transportation planning, including multinomial and nested logit models.

## Acknowledgments

- Webinar prepared by Cambridge Systematics, Inc.
  - Thomas Rossi
  - Yasasvi Popuri
- Prepared for Federal Highway Administration
  - H. Sarah Sun
- Arranged by Texas Transportation Institute
  - Gary Thomas

2

### Key Message: Personnel Behind the Webinar Series

#### Details:

**FHWA Manager:** H. Sarah Sun conceived and designed the overall structure and content of this Webinar series in conjunction with Cambridge Systematics, Inc.

**Presenters: Thomas Rossi and Yasasvi Popuri, Cambridge Systematics, Inc.**

**Thomas Rossi** is a Principal of Cambridge Systematics with 25 years of experience in transportation planning and travel demand forecasting. He has developed and applied trip based and activity based models throughout the U.S. For the past 15 years, Tom has been a consultant to USDOT for model improvement research and development/teaching of training courses. He is the Chairman of TRB Committee on Transportation Demand Forecasting. Tom holds Bachelor's degrees in Civil Engineering and mathematics and a Master's degree in Transportation from the Massachusetts Institute of Technology.

**Yasasvi D. Popuri** is a Senior Associate of Cambridge Systematics with experience in the fields of travel demand forecasting, discrete choice modeling, market research analysis, geographical information systems (GIS), and transportation planning. He has managed projects involving transit ridership forecasting, travel model estimation, and freight modeling. Mr. Popuri received a Master's degree in Transportation Systems Engineering from the University of Texas at Austin, and a Bachelor's degree of Technology from the Indian Institute of Technology at Madras.

## Webinar Objectives Learning

- Intended for those who have a low level of familiarity with the estimation and validation of travel models
- Introduces
  - Development of model estimation data sets
  - Structures of various model components
  - Procedures for estimating models
- Does not train participants in model estimation procedures

3

### **Key Message: Content and Objectives of Webinar 1**

#### **Details:**

At the end of the webinar, participants will be able to better manage model development done for them by others (e.g., consultants) and to understand and evaluate the results they are presented with.

Participants will understand the key steps in a typical travel demand model. They will have the tools to determine the quality of the models presented to them and to evaluate the capabilities of the model.

## Webinar Outline

- Session 1: Introduction – October 16, 2008
- Session 2: Data Set Preparation – November 6, 2008
- Session 3: Estimation of Non-Logit Models – December 11, 2008
- Session 4: Estimation of Logit Models – February 10, 2009

4

### **Key Message: Timeline of the Webinar Series**

#### **Details:**

The first session on October 16, 2008 will be followed by a second session on November 6, 2008 and a third session on December 11, 2008.

The timeline for all the subsequent webinars (4 through 8) is yet to be determined. An announcement email will be sent once these dates are finalized.

## Webinar Outline (continued)

- Session 5: Application and Validation of Logit Models – March 12, 2009
- Session 6: Advanced Topics in Discrete Choice Models – April 14, 2009
- Session 7: Trip Assignment – May 7, 2009
- Session 8: Evaluation of Validation Results – June 9, 2009

5

### **Key Message: Timeline of the Webinar Series**

#### **Details:**

The first session on October 16, 2008 will be followed by a second session on November 6, 2008 and a third session on December 11, 2008.

The timeline for all the subsequent webinars (4 through 8) is yet to be determined. An announcement email will be sent once these dates are finalized.

## Homework

(You thought you finished this forever decades ago!)

- Exercise will be assigned at the end of each session, based on some of the information discussed
- May require generally available software (e.g., Excel, text editor, etc.), but not specialized software such as model estimation or application programs
- You may e-mail questions on the homework to us up to one week before the next session
  - [trossi@camsys.com](mailto:trossi@camsys.com), [yopuri@camsys.com](mailto:yopuri@camsys.com)
- Homework will be reviewed during the first part of the next session

6

### Key Message: Homework Details

#### Details:

The webinar will reinforce the concepts discussed during the session using homework assignments. No special software other than Microsoft Excel and Word is required to solve the homework.

Participants may email any questions regarding the homework to the presenters. A small portion of each subsequent session will be dedicated to discussing the solution and also addressing any questions and comments that participants may have about the solutions.

## Types of Urban Models

- Four-step models
- Three-step models (no mode choice)
- Tour and activity based models

7

### **Key Message: Urban Model Types**

#### **Details:**

Urban models can be broadly classified into three categories:

1. The most common models are the four-step models, which will be discussed in detail in this webinar. The four most important steps are Trip Generation, Trip Distribution, Mode Choice and Trip Assignment.
2. In smaller urban areas with very little public transit service, the mode choice component is not very useful and is often not included.
3. More recently, tour- and activity-based models are gaining popularity. These models are more advanced and more behaviorally realistic than the other two types of models.

## Model Components Four-Step Models

- Trip production
- Trip attraction
- Trip distribution
- Mode choice
- Assignment
- Time of day
- Auto ownership
- Other
  - Trucks/freight
  - External trips
  - Other?

8

### Key Message: Contents of a Four-Step Model

#### Details:

The four most important steps of a four-step model are:

1. **Trip Generation:** This step addresses the question of how many trips are produced from each zone and how many are attracted to each zone in the study area. The production is the home end of a trip and the attraction end is the non home-end of a trip. Note that the production and attraction are not the same as origin and destination ends. Trips with neither end at home are called non home-based trips. For the non home-based trips, the production end is defined to be the same as the origin end, the attraction end to be the same as the destination end. Essentially, the trip generation step gives us the row and column totals of the trip table.
2. **Trip Distribution:** This step tells us where the trips go. In other words, given the row and column totals (from trip generation), the proximity measures between each pair of zones, and the attractiveness of each zone, trip distribution tells us how to populate the individual cells of the matrix.
3. **Mode Choice:** Mode choice provides a framework to apportion the total trips among various transportation modes such as drive alone, shared ride, transit, walk, bike etc.
4. **Assignment:** Trip assignment uses the trip tables along with the highway and transit networks and generates traffic volumes and transit ridership.

In addition to these four common models, many other components are often present within the model framework. These include time of day models to further apportion trips tables by time period, auto ownership models which provide a capability to predict and forecast the number of cars a household is likely to own given its demographic and socio-economic makeup and its location.

More details are presented in FHWA-NHI-152054 Introduction to Urban Travel Demand Forecasting course.



## The Model ...

- ... takes a set of *input data* ...
- ... and converts it to a set of *output data* ...
- ... using a set of *mathematical models* ...
- ... which use *parameters* to perform the conversions
- The input and output data for individual model components may be temporary, or interim

9

### Key Message: What is a Model?

#### Details:

Travel models are **tools** to create data (**outputs** of the travel model) needed by planners to assist in planning analyses. They are created from other available data (the model **input** data) using a set of **mathematical models**. These model have parameters that must be estimated (or otherwise obtained) and validated.

The input and output data may be temporary, being created by one model component mainly for input to another component. Examples include travel time skim matrices, trip ends, and trip tables.

## Model Parameters ...

... may be:

- Estimated (usually from local data),
- “Borrowed” from another model or other data source, or
- Asserted, based on knowledge from other sources

10

### **Key Message: What are Model Parameters?**

#### **Details:**

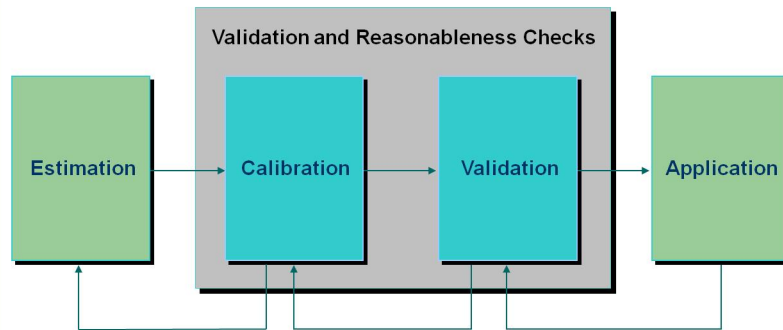
What are some examples of model parameters?

There are many examples:

1. Trip generation rates
2. Friction factors for gravity model
3. K-factors for gravity model
4. Mode choice utility coefficients
5. Time of day factors
6. “A” and “B” in BPR volume-delay function
7. And many others

Model estimation is concerned with obtaining these parameters using household survey data or other sources such as onboard data, census journey-to-work data etc.

## The Model Development Process



**Key Contents: What does the model development process comprise?**

**Details:**

In model estimation, the parameters are developed. Commonly, parameters are estimated using statistical techniques using data from surveys or other sources.

The validation of the model involves checks of the parameters and the results, compared to other data or results from other model runs. Any issues identified during validation may result in adjustments to parameters (calibration). When parameters have been changed, the model must be revalidated.

The results of validation could involve re-estimation of parameters in some cases. For example, if the validation process identified that, say, mode choice results were correct overall but not by auto ownership level, the model might be reestimated with a new or revised variable dealing with auto ownership.

Model application can occur once the model has been validated for the base year. However, application for the forecast year is a necessary part of model validation and can lead to further calibration changes.

## Model Development Process

- Have a model development plan
- Assemble necessary data for estimation, validation, and application
- For each component
  - Define the mathematical model and the model structure
  - Estimate (or otherwise obtain) the parameters
  - Validate the model (a whole 'nother topic!)
  - Recalibrate the parameters as necessary

12

### Key Contents: Model Development Process

#### Details:

The model development process begins well before any model is actually estimated. A significant effort must be spent upfront to do the following:

1. Identify the key purposes for which the model will be used. These include both the primary and secondary purposes for which the model is expected to be used.
2. Identify key sources of data such as household surveys, onboard surveys, census data, employment data, land use estimates, zones and network structure.
3. Procure the data sources in advance and make sure that good documentation is available.

Once these basic steps are undertaken, the estimation process begins. This includes the following steps:

1. Define the structure of each mathematical model and the variables that will be used.
2. Estimate or otherwise obtain the parameters.
3. Apply the model and validate its performance using available data sources such as household surveys, census data, onboard data, traffic counts etc.
4. If the model is not validated, then identify ways to recalibrate the model parameters.

## Model Development Process (continued)

- Validate overall model system, including forecast years
  - Recalibrate as necessary
- Don't forget documentation!

13

### Key Contents: Model Development Process

#### Details:

Note that forecast years cannot be validated using any observed data. However, the model can be checked for its sensitivity to various ranges of land use, socio-economic, highway and transit network forecasts. If past versions of model exist, a trend analysis can also be done to check the reasonableness of model forecasts.

Detailed documentation of the model is extremely important not only for guiding current users but also for providing a solid background for any future updates. It also adds transparency to the model development process and thus the credibility to the model.

## Model Application Process

- Overall inputs
  - Socioeconomic data
  - Networks (highway, transit)
  - Other (parking costs, auto operating costs, etc.)
- Overall outputs
  - Link volumes → VMT
  - Link speeds → VHT
  - Transit boardings, line volumes
  - Trips by O-D, mode, purpose, time of day, etc.

14

### Key Contents: Model Application Process

#### Details:

Key inputs to the transportation model include:

1. Zonal socioeconomic data such as:
  - Number of households classified by their size, vehicle ownership, number of workers.
  - Population
  - Employment, typically classified into retail, non-retail, commercial, industrial
2. Networks
  - Highway networks
  - Transit networks
3. Other
  - Parking costs
  - Auto operating costs

Key outputs from the model include:

1. Trip tables by mode, purpose and time of day.
2. Traffic volumes, speeds, VMT and VHT on each link in the model network by time of day.
3. Transit boardings on all the transit routes by time of day.

## Some Model Component Examples

- Trip generation
  - Input – Socioeconomic data
  - Output – Trip ends
- Trip distribution
  - Input – Trip ends/network skims
  - Output – Trip tables

15

### Key Contents: Model Component Examples

#### Details:

##### Trip Generation:

This step addresses the question of **how many** trips are produced from each zone and how many are attracted to each zone in the study area. The **production** is the home end of a trip and the **attraction** end is the non home-end of a trip. Note that the production and attraction are not the same as origin and destination ends. Trips with neither end at home are called non home-based trips. For the non home-based trips, the production end is defined to be the same as the origin end, the attraction end to be the same as the destination end. Essentially, the trip generation step gives us the row and column totals of the trip table.

Inputs to the trip generation step include:

1. Socio-economic data such as number of households of various sizes, vehicle ownership levels, number of workers etc
2. Employment and population in each zone.
3. Trip production rates for each type of household generated from the household survey,
4. Trip attraction rates generated from the household surveys

Outputs from the trip generation step include:

1. Number of trip ends produced from each zone
2. Number of trip ends attracted to each zone.

##### Trip Distribution:

This step tells us **where** the trips go. In other words, given the row and column totals (from trip generation), the proximity measures between each pair of zones, and the attractiveness of each zone, trip distribution tells us how to populate the individual cells of the matrix.

Inputs to the trip distribution step include trip ends from the trip generation step and the network skims that will be used to determine impedances between various pairs of zones.

Outputs include person trip tables by trip purpose.

## Some Model Component Examples (continued)

- Mode choice
  - Input – Trip tables/network skims
  - Output – Trip tables by mode
- Highway assignment
  - Input – Trip tables for auto mode/highway network
  - Outputs – Volumes/speeds

16

### Key Contents: Model Component Examples

#### Details:

#### Mode Choice:

Mode choice provides a framework to apportion the total trips among various transportation modes such as drive alone, shared ride, transit, walk, bike etc.

Key inputs to the mode choice model include:

1. Transit Modes:
  - In-Vehicle Travel Time, Boarding times, Wait times, Walk times
  - Transit fares
  - Parking costs at park and ride stations
2. Highway Modes:
  - Highway travel times and distances
  - Operating costs
3. Non-Motorized Modes:
  - Travel times by walk/bike using walk/bike paths
4. Total person trip tables from trip distribution.

Outputs from the mode choice model are trip tables by each mode.

#### Highway Assignment:

The highway assignment step uses the time period-specific highway trip tables obtained from the mode choice step and assigns the trips to the network. Key inputs to the highway assignment step include:

1. Trip tables for auto modes
2. Properly coded highway network with link directionalities and capacities.

Outputs from the highway assignment step include volumes on each link in the model network and the speeds.



## Common Model Component Formulations

- Simple factoring
- Cross-classification
- Regression
- Logit (multinomial, nested)
- Assignment

17

### Key Contents: Common Model Formulations

#### Details:

Some of the most common types of mathematical formulations used in travel demand models are as follows:

1. Simple factoring
2. Cross-Classification
3. Regression
4. Multinomial and Nested Logit Models
5. Shortest path/capacity-constrained algorithms for assignment.

Simple factoring – the model inputs at some level are multiplied by a set of factors (the model parameters) to produce the outputs. Example: time of day factors applied to daily trips to get trips by time period.

Other model types described on subsequent slides.

## Model Estimation Data Sources

- Household activity/travel survey (household, trip level)
- Transit on-board survey
- Critical nonsurvey data
  - Socioeconomic data
  - Networks
  - Other (area types, parking costs, auto operating costs, etc.)

18

### Key Contents: Model Estimation Data Sources

#### Details:

For models that are estimated from local data, the main data source is the household activity/travel survey. The data can be used to estimate parameters for trip generation, trip distribution, mode choice, time of day, and vehicle availability models, among others.

Transit on-board survey data, while also the most important data source for information on current transit ridership and validation of transit related model components, may also provide information for mode choice model estimation, since transit trip records may be sparse in household survey data.

Socioeconomic data are needed for models with variables involving demographic characteristics or densities. Network skim data are needed for models with transportation level of service variables (such as trip distribution and mode choice). Data on any other variables that will appear in the models are also needed.

The importance of data quality cannot be overemphasized! Poor data result in poor models.

#### Questions for Consideration:

What other data might be used?

## Model Types

Model Component	Typical Model Type	Usual Estimation Data Source	Other Basic Data
Auto ownership	Multinomial/ordered response logit	Household survey (household file)	Socioeconomic, network (skims)
Trip production	Cross-classification	Household survey (household file)	Socioeconomic
Trip attraction	Linear regression	Household survey (trip file)	Socioeconomic
Trip distribution	Gravity/multinomial logit	Household survey (trip file)	Socioeconomic, network (skims)

19

### Key Contents: Model Estimation Data Sources

#### Details:

This slide presents some of the most common model types used for auto-ownership models, trip production, trip attraction and trip distribution models.

Also presented here are the common estimation data sources and other basic data used. Please refer to slides 15 and 16 for details on the data items needed for estimation of each of these models.

## Model Types (continued)

Model Component	Typical Model Type	Usual Estimation Data Source	Other Basic Data
Time of day	Simple factoring	Household survey (trip file)	
Mode choice	Multinomial/nested logit	Household survey (trip file)	Socioeconomic , network (skims)
Highway assignment	Static user equilibrium		Highway network
Transit assignment	All-or-nothing, multipath		Transit network

20

### Key Contents: Model Estimation Data Sources

#### Details:

This slide presents some of the most common model types used for time-of-day, mode choice, highway assignment and transit assignment models.

Time-of-day models most often consist of simple factoring techniques. The household survey data are used to obtain the proportion of daily trips that happen in each major time period. These proportions or factors are then applied to the overall daily trip table to get trip tables for each time period. Advanced formulations of time-of-day models have also been used. These include logit models for estimating the split of total trips for each purpose.

Mode choice models adopt multinomial or nested logit models that are estimated using the household survey data and the network levels of service.

Highway assignment models typically use static user-equilibrium models, while transit assignment models use all-or-nothing or multipath assignments.

## Regression Model

$$Y = B_0 + B_1 X_1 + B_2 X_2 + \dots + B_n X_n$$

where:

**Y** = Dependent variable

**B<sub>i</sub>** = Estimated coefficients

**X<sub>i</sub>** = Independent variables

21

### Key Contents: Regression Model

#### Details:

Regression models are the most commonly used modeling techniques. Regression models relate a dependent variable (Y in the slide above) to a bunch of independent variables (the X's in the slide above). Thus given the value of the X's, one would be able to predict a value for Y.

The B's in the slide above represent "parameters" or the "coefficients" of the regression. When one talks about "running" a regression, one is referring to the estimation of the B's in the parameters.

On this and the following slides, the model parameters are shown in red.

## Regression Model Example Trip Attraction Model

HBO Attractions = **1.32** (service employment)  
+ **1.46** (retail employment)  
+ **0.76** (households)

22

### Key Contents: Regression Model – An Example

#### Details:

This slide presents an example of a sample regression model. Here are all the details of this regression:

The dependent variable is the number of trip ends attracted to each model zone.

The independent variables are service and retail employments and the number of households in each zone.

The way to interpret this equation is as follows: each additional unit of service employment (each job in the service sector) will create 1.32 additional trips to the zone, everything else remaining equal. Similarly, each additional retail employee will create 1.46 additional trips and each additional household will attract 0.76 additional trip to the zone, everything else being equal.

## Cross-Classification Model

Independent Variable #2	Independent Variable #1				
		Value 1	Value 2	...	Value n
	Value 1	Dep var value	Dep var value		Dep var value
	Value 2	Dep var value	Dep var value		Dep var value
	...				
	Value n	Dep var value	Dep var value		Dep var value

23

### Key Contents: Cross-Classification Models

#### Details:

Cross-classification is the most commonly used technique for trip production models. Cross-classification involves the following items:

1. Identify independent variables that will be used to classify all households. Most common variables include household size and vehicles. The rationale is that households of a given size and with a given number of vehicles have similar necessity to make out-of-home trips. For example, all households with four people and two vehicles in the household are very similar in the number of trips they make each day. The higher the household size, the higher the necessity to make trips on an average and the higher the number of vehicles, the higher the number of vehicles, the higher the number of trips made. If there are 4 categories of household size and 4 categories of vehicle ownership, we have a total of 16 cross-class cells. That is, the matrix above will have 16 cells.
2. Identify the number of households that fall in each category. This is easily obtained from household surveys. For example, one can easily find how many households in the study region have four members in the household and 2 vehicles in the household. Similarly, the number of households that fall in each of the 16 cells discussed above can be obtained and tabulated.
3. Identify trips of each purpose made by the households of each type. That is, corresponding to each of the 16 cells above, one can find how many trips of each purpose are made. This can once again be obtained from household travel surveys.
4. Once we have the total trips of each type and the underlying number of households, we can find the trip rate for each purpose.
5. The outcome of this process is quite simply a set of production rate tables, one for each purpose. The cells in each table represent the average number of trips made by each household characterized by that cell.

The purpose of deriving these cross-class rates is that if one were able to forecast the number of households that fall in each category, one would also be able to forecast the number of trips that these households make by multiplying the total number of households with the trip rates.

## Cross-Classification Model Example

### Trip Production Model – Home Based Shop Trips

Persons/ Household	Vehicles/Household				
		0	1	2	3+
	1	0.294	0.333	0.333	0.333
	2	0.377	0.465	0.619	0.619
	3	0.425	0.515	0.619	0.619
	4+	0.433	0.601	0.740	0.767

24

#### Key Contents: Cross-Classification Models – An Example

##### Details:

This table illustrates the description provided in the previous slide. There are four categories for household size (1,2,3,4+) and four categories for vehicle ownership (0,1,2,3+). Thus, there are 16 household categories.

The table above indicates that each household with 3 persons and 1 vehicle makes 0.515 daily home-based shopping trips on an average.



## Gravity Model Trip Distribution

$$T_{ij} = \frac{P_i A_j F(t)_{ij} K_{ij}}{\sum_j P_i A_j F(t)_{ij} K_{ij}}$$

where:

- $T_{ij}$  = number of trips produced in zone i and attracted to zone j
- $P_i$  = trips produced in zone i
- $A_j$  = trips attracted to zone j
- $F(t)_{ij}$  = friction factor from i to j (based on impedance t)
- $K_{ij}$  = K factor from i to j
- i = origin zone
- j = destination zone

25

### Key Contents: Gravity Model – Trip Distribution

#### Details:

The gravity model simply gives a way of calculating the trip interchanges, given the productions and attractions.

Note that in the above slide, the  $P_i$  and  $A_j$  are the trips produced from zone i and  $A_j$  represents trips attracted to zone j.

Given these two numbers and the proximity of zones i and j, the gravity model just tells us how many trips produced from zone i are attracted to zone j.

The K-Factors are simply used to adjust for items other than proximity and impedance that influence distribution such as area type of the zone (CBD, urban etc.)

## Multinomial Logit Model

$$P(\text{alt } 1) = \frac{\exp(V_1)}{\sum_j \exp(V_j)}$$

where:

$V_j$  = Deterministic component of utility of alternative  $j$   
exp = exponential function ( $e^x$ )

26

### Key Contents: Multinomial Logit Model

#### Details:

The multinomial logit model is a powerful statistical tool to model discrete individual choices. A prominent application of the logit model is the mode choice model. A mode choice model tries to predict the probability that an individual with a given set of attributes chooses a particular transportation mode. The attributes of a given mode are captured by the utility of the mode.

The logit model formulation states that if there are three modes auto, transit and walk, then the probability that an individual chooses auto is the ratio of the exponentiated utility of the auto mode to the sum of the exponentiated utilities of auto, transit and walk modes.

The interesting point to note is that the probability of a mode is dependent not only on its attributes but also on the attributes of its competing modes.

## Multinomial Logit Model Utility Function

$$V_{ij} = B_{0j} + B_{1j} X_{1ij} + B_{2j} X_{2ij} + \dots + B_{nj} X_{nij}$$

where:

$V_{ij}$  = Utility (deterministic component) of alternative j for individual i

$X_{kij}$  = Attributes ( $k = 1, n$ ) for individual i for alternative j

$B_{kj}$  = Estimated coefficients for attribute k for alternative j

$B_{0j}$  = Alternative-specific constant for alternative j

27

### Key Contents: Multinomial Logit Model – Utility Function

#### Details:

So what does the utility function contain? We already said that the utility function is a representation of the attributes of a particular alternative. Going back to the mode choice example, for the auto mode, the utility function will simply be a linear combination of auto travel times, operating costs, and the person's own socio-economic attributes. As with the regression models, the estimation of a mode choice model focuses on obtaining the parameters or the coefficients (B's) in the equation above.

## Multinomial Logit Model Example Vehicle Availability Model

Variable	Vehicle Availability Level			
	0	1	2	3+
Constant	Base alt (U = 0)	0.64	-0.45	-2.29
1 worker in household (0,1)		0.83	1.10	1.66
2+ workers in household (0,1)		0.54	2.47	3.32
Low-med income (0,1)		1.16	2.18	2.26
High-med income (0,1)		0.87	3.04	3.64
High income (0,1)		1.78	4.31	5.28
% employment within 15 min		-0.03	-0.08	-0.12

28

### Key Contents: Multinomial Logit Model – Example

#### Details:

Here we present an example of the vehicle availability model. As with mode choice, vehicle availability models are also logit models that model the probability of a household owning a given number of vehicles.

Let us interpret the results shown in the table above. To do this, we first translate the tabular results above into individual utility expressions. Obviously, there will be as many utility functions as there are alternatives. So, in this case, we will need to write out 4 utility expressions. The utility expressions and their interpretation are shown in the next page.

## Multinomial Logit Model Example Vehicle Availability Model (continued)

Utility functions:

$$U_0 = 0$$

$$U_1 = 0.64 + 0.83 (1 \text{ worker}) + 0.54 (2+ \text{ worker}) \\ + 1.16 (\text{Low-medium income}) + 0.87 (\text{High-medium income}) \\ + 1.78 (\text{High income}) -0.03 (\% \text{ employer w/in 15 min})$$

$$U_2 = -0.45 + 1.10 (1 \text{ worker}) + 2.47 (2+ \text{ worker}) \\ + 2.18 (\text{Low-medium income}) + 3.04 (\text{High-medium income}) \\ + 4.31 (\text{High income}) -0.08 (\% \text{ employer w/in 15 min})$$

$$U_3 = -2.29 + 1.66 (1 \text{ worker}) + 3.32 (2+ \text{ worker}) \\ + 2.26 (\text{Low-medium income}) + 3.64 (\text{High-medium income}) \\ + 5.28 (\text{High income}) -0.12 (\% \text{ employer w/in 15 min})$$

29

**Errata:** Please note that the last utilities are all titled U1. These must be U1, U2 and U3.

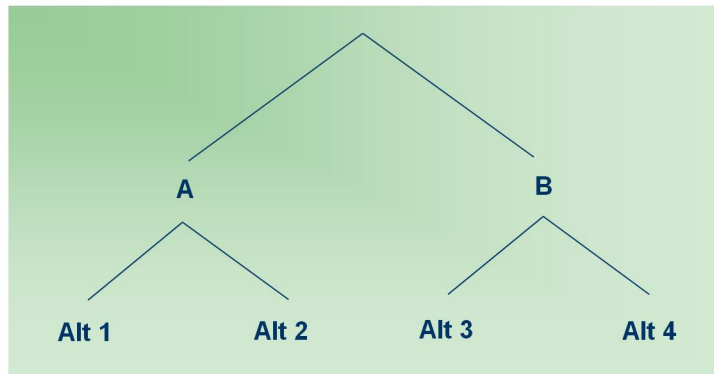
### Key Contents: Multinomial Logit Model – Example

#### Details:

So how do we interpret the results?

1. First, we notice that the zero-vehicle alternative has a utility of zero. This is because in logit models, we are interested only in the relative magnitudes of utilities and not the absolute magnitudes. In fact, we can never estimate the absolute utility of each alternative for each individual. So, we set the utility of one of the alternatives as zero. This simplifies the interpretation of the parameters or the coefficients.
2. Second, we notice that the constant term is positive for the 1-vehicle alternative and becomes more and more negative as we move to higher ownership levels. This indicates that everything else remaining equal, households are more likely to own one vehicle.
3. Third, we also notice that the worker and income variables become increasingly positive as we move from 0 and 1-vehicle to 2- and 3+-vehicle households. This is intuitive because the higher the number of workers in a household the higher the likelihood of owning 2 or 3 vehicles. The same logic applies to the income variables as well.
4. Finally, we notice that the variable called “percentage of employment within 15 minutes of highway travel time” has a negative coefficient. This is because the higher the employment within a short travel distance, the lesser the need for a vehicle to access jobs! This is why we notice an increasing negative trend as we move from 1 –vehicle to 3+-vehicle households.

## Nested Logit Model



30

### Key Contents: Nested Logit Model – Example

#### Details:

The multinomial logit model treats every single alternative as being completely independent of the other alternatives.

The nested logit model recognizes the fact that some alternatives are in reality more like each other than the others. So in the figure above, alternatives 1 and 2 are siblings, which share common unobserved attributes.

A good example of nested logit models would be mode choice models again! Our previous examples talked about transit as a single alternative. But most often, we have many transit modes like rail and bus. Besides, there are multiple access modes like walk and drive. So how can we represent all these individual alternatives? One way is to represent them in a multinomial model form. So, bus with walk access, bus with drive access, rail with walk access and rail with drive access will all be assumed to be independent alternatives. This does not recognize the similarities within the bus modes or within the rail modes. So one possible improvement could be to have a nested logit model that has the bus and rail modes separated out, and the walk and drive access sub-modes as siblings under the bus or rail nest.

## Nested Logit Model

$$P(\text{alt } 1 | A) = \frac{\exp(V_1)}{\exp(V_1) + \exp(V_2)}$$

$$P(A) = \frac{\exp[f(V_1, V_2)]}{\exp[f(V_1, V_2)] + \exp[f(V_3, V_4)]}$$

$$P(\text{alt } 1) = P(\text{alt } 1 | A) P(A)$$

31

### Key Contents: Nested Logit Model – Example

#### Details:

The nested logit models are a generalization of the multinomial logit model. The probability expressions are shown above. The expressions above simply indicate that the probability of an individual choosing alternative 1 is the product of two probabilities: the probability of the individual choosing the nest A (the  $P(A)$  in the slide above) and the conditional probability of the individual choosing alternative 1 given that he or she chose nest A (the  $P(\text{alt } 1|A)$  in the slide above)

## Highway Assignment

- Static equilibrium assignment
  - Assumes O-D travel times along all used paths are equal
- Link travel time is a function of travel time, e.g.,
$$T = T_0 + [ 1 + a (v/c)^b ]$$

where:

$T$  = link travel time

$T_0$  = free flow travel time

$v$  = link volume

$c$  = link capacity

$a, b$  model parameters

32

### Key Contents: Highway Assignment

#### Details:

Session 7 will discuss the assignment process and the development of parameters and models such as volume-delay functions.

It should be noted that for a complete discussion of the implementation of the trip assignment process, the NHI course on Introduction to Urban Travel Demand Models provides a significant amount of information.



## Model Parameter Development Example

33

- Estimated (from household survey)
  - Trip generation rates
  - Friction factors for gravity model
  - Time of day percentages by purpose
- Borrowed
  - Mode choice utility coefficients (from other model)
- Asserted
  - BPR function parameters a, b (from literature)
  - K-factors (all set to 1.0)

### Key Contents: Model Parameter Development Example

#### Details:

The information on this slide is an example of how model parameters may be estimated, borrowed, and asserted within a single model. This may differ for individual models. For example, the parameters shown as borrowed or asserted might be estimated in some models.